Identical Consensus Sequence and Conserved Genomic Polymorphism of Hepatitis E Virus during Controlled Interspecies Transmission

Jerome Bouquet,[a,b,c] Justine Cheval,[d] Sophie Rogée,[a,b,c] Nicole Pavio,[a,b,c] and Marc Eloit[a,b,c,d,e]

UMR 1161 Virology, ANSES, Laboratoire de Santé Animale, Maisons-Alfort, France[a]; UMR 1161 Virology, INRA, Maisons-Alfort, France[b]; UMR 1161 Virology, Ecole Nationale Vétérinaire d'Alfort, Maisons-Alfort, France[c]; Pathoquest, Paris, France[d]; and Department of Virology, Institut Pasteur, Paris, France[e]

High-throughput sequencing of bile and feces from two pigs experimentally infected with human hepatitis E virus (HEV) of genotype 3f revealed the same full-length consensus sequence as in the human sample. Twenty-nine percent of polymorphic sites found in HEV from the human sample were conserved throughout the infection of the heterologous host. The interspecies transmission of HEV quasispecies is the result of a genomic negative-selection pressure on random mutations which can be deleterious to the viral population. HEV intrahost nucleotide diversity was found to be in the lower range of other human RNA viruses but correlated with values found for zoonotic viruses. HEV transmission between humans and pigs does not seem to be modulated by host-specific mutations, suggesting that adaptation is mainly regulated by ecological drivers.

Hepatitis E virus (HEV) is a causative agent of acute hepatitis in humans. The disease is usually self-limited but is a major public health concern both in developing countries, where it causes large waterborne epidemics, and in industrialized countries, where sporadic autochthonous cases of unclear origin are reported. It is the only hepatitis virus that infects animals other than primates, such as swine, wild boars, and deer (28). Direct zoonotic transmissions through consumption of contaminated food were observed in a few cases in a region where HEV is not endemic (20, 38).

HEV is a positive single-stranded RNA virus. It is the sole member of the *Hepeviridae* family and the *Hepevirus* genus (25). HEV isolates have been divided into at least four genotypes, two putative genotypes, and 24 subtypes (12, 16, 22, 25). Genotype 1 and 2 are present in humans only, while genotypes 3 and 4 can infect both humans and animals (28). The 7.2-kb genome of HEV is composed of three open reading frames (ORF). ORF1 encodes a nonstructural polyprotein with six conserved domains and one hypervariable region (14, 19). ORF2 encodes the capsid protein, and ORF3 encodes a phosphoprotein necessary for infection *in vivo* (9).

Many RNA viruses circulate as a population of heterogeneous but closely related genomes within the same individual. The emergence of such quasispecies is the consequence of a high mutation rate engendered by the activity of nonproofreading RNA-dependent RNA polymerases (RdRP), coupled with a high replication rate.

To date, the only description of the HEV quasispecies has been obtained by restriction fragment length polymorphism and sequencing of a 448-bp fragment from HEV genotype 1 (10). Since genotype 1 is restricted to humans, the results of this study had limited significance for HEV population variability in other host species and for the existence of a putative species barrier for genotypes 3 and 4 (22). Swine isolates of genotype 3 and 4 HEV can infect primates, and human isolates of genotype 3 and 4 HEV have been shown to replicate in pigs (1, 8, 11, 23, 24). The objective of the present study was to analyze the genomic diversity of full-length HEV genotype 3 during a single passage between the two different host species

using high-throughput sequencing (HTS) techniques and deep genomic variability analysis.

MATERIALS AND METHODS

Experimental infection. Human fecal samples were collected from a French patient with no recent travel history outside France who had developed an acute autochthonous hepatitis E of subtype 3f, according to the classification of Lu et al. (22).

Two 3-month-old pigs were orally inoculated using industrially sterilized pet food mixed with 1 g of human sample infected with 2×10^9 copies of HEV RNA. After inoculation, feces of pigs were collected every 2 days for 1 month, and bile samples were collected after a light surgical procedure at 15 days postinfection (dpi) (Fig. 1). This experimental protocol was validated by the ethics committee (ComEth; saisine number 10-0041) from the National Veterinary School of Alfort, the National Agency for Safety, and University Paris 12.

Serological analyses. Serological analyses were conducted as previously described (32). Briefly, serum samples were tested with an anti-HEV total immunoglobulin kit for human diagnosis (EIAgen HEV Ab Kit; Adaltis, Ingen, France), replacing the secondary antibody by a peroxidase-conjugated rabbit polyclonal anti-pig IgG(H+L) (Abcam, France). Samples were considered positive when the optical density at 450 nm (OD_{450}) ratio of the sample to the cutoff value (equal to the value of the negative control + 0.350) was >1.

RNA extraction and quantitative RT-PCR. HEV load was estimated by real-time reverse transcription-PCR (RT-PCR). Total RNA was extracted from 200 μl of fecal samples in 10% phosphate-buffered saline (PBS) or bile samples using a viral QiAmp kit (Qiagen, Courtaboeuf, France) according to the manufacturer's instructions, and real-time RT-PCR, as developed by Jothikumar et al., was performed on 2 μl of RNA using a Quantitect RT-PCR probe (Qiagen, Courtaboeuf, France) (17). A LightCycler apparatus (Roche Molecular Biochemicals, Meylan, France) was used for sample analysis. Standard quantification curves were calculated with standard HEV RNA of subtype 3f. The standard plasmid was
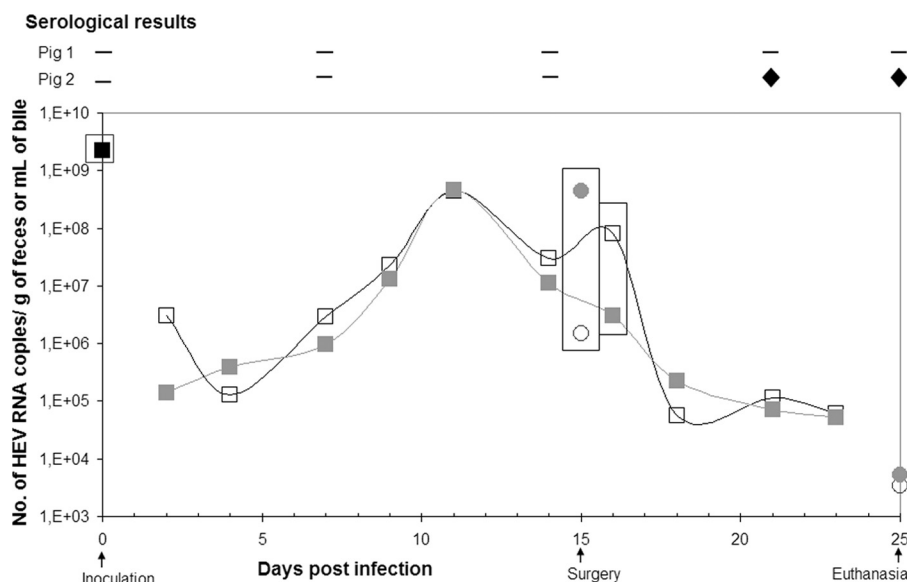
Serological results

FIG 1 Experimental infection of two pigs with human HEV. The HEV load of initial human fecal sample inoculated to pigs is plotted as a black square. Excretion of the virus in the feces of pigs is plotted as white squares for pig 1 and as gray squares for pig 2. The HEV load in the bile is plotted as white dots for pig 1 and as gray dots for pig 2. The presence of anti-HEV IgG in serum is indicated by black diamonds. Samples used in high-throughput sequencing are boxed in black. Times of inoculation, surgery, and euthanasia are indicated on the axis for days postinfection.

constructed by cloning a fragment corresponding to the genomic region from nucleotides (nt) 5190 to 5489 of a French swine HEV sequence of genotype 3f (accession number JF718793) into the NheI/XhoI-digested pCDNA 3.1 (Life Technologies, Villebon sur Yvette, France) Amplification and cloning were performed using forward (5′-NheI-CTGCATCGC CCATGGGATCGC-3′) and reverse (5′-XhoI-CGCTGGGACTGGTCAC GCC-3′) primers.

The HEV-positive human fecal sample, a pool of swine fecal samples collected at 16 dpi, and a pool of swine bile samples collected at 15 dpi were subjected to HTS.

Sequence-independent amplification. After DNase treatment for 2 h at 37°C (0.33 U/μl of sample; Qiagen, France), total nucleic acids were extracted using a Nucleospin RNA virus kit (Macherey-Nagel, Germany) and then amplified without use of HEV-specific PCR primers, as described previously (6). Briefly, bacteriophage Φ29 polymerase-based multiple-displacement amplification was preceded by a cDNA synthesis step performed with random hexamer primers. Ligation and whole-genome amplification were then performed with a QuantiTect whole-transcriptome kit (Qiagen, France) according to the manufacturer's instructions.

High-throughput sequencing. Illumina GAII sequencing was subcontracted to GATC (Constance, Germany). High-molecular-weight DNA (5 g), resulting from genomic RNAs as described above, was fragmented into 200- to 350-nt fragments, to which adapters were ligated. Adapters included a nucleotide tag allowing for multiplexing of the three samples in one channel.

Data filtering and establishing consensus sequences. Illumina sequencing data were processed by using a bioinformatic analysis pipeline as described previously (6). Briefly, Illumina sequence reads were trimmed of their low-quality score extremities, and host genome sequences (*Homo sapiens* and *Sus scrofa*) scanned with SOAPaligner (http://soap.genomics .org) were discarded. A quick and very restrictive BLASTN study was also performed to eliminate additional host reads. BLASTN and BLASTX were used to scan dedicated specialized viral, bacterial, and generalist databases maintained locally (GenBank viral and bacterial databases) (6). Reads and contigs matching HEV sequence were mapped over the closest sequence hit using relaxed alignment settings (length fraction, 0.5; similarity, 0.8) in the CLC Genomics Workbench (CLC bio, Cambridge, MA).

Validation of polymorphic sites and analysis of sequence diversity. To eliminate overmutated reads generated by the technique, a new mapping of reads matching HEV sequences was performed with SOAPaligner (http://soap.genomics.org) on the newly assembled HEV consensus sequences, removing reads with more than two mismatches.

The error rate due to the amplification and sequencing processes was established by observing the variability of conserved genes from host species and bacteria present in the samples following the same filtering process as HEV sequences. The number of sequencing errors was plotted against the number of nucleotides mapped over a consensus sequence to uncover the error rate. A theoretical number of mutations generated by the technique was calculated for each nucleotide position of HEV quasispecies by multiplying the error rate with the coverage at each position and rounded to the immediate upper whole number. Polymorphic sites were validated when the observed number of a base (or gap) different from the consensus sequence was superior to the theoretical number of mutation errors.

Polymorphism parameter calculation. To define the intrahost diversity of HEV quasispecies, the genome-wide data of validated nucleotide sites were analyzed to measure the average nucleotide diversity and mean diversity. Nucleotide diversity as developed by Nei and Li (27) was calculated as the average percentage of single nucleotide polymorphism (SNP) over the genome, whereas mean diversity corresponds to the percentage of the number of substitutions divided by the total number of nucleotides.

For each sample, a theoretical sequence containing all validated mutations was created. Selective pressure along the three ORFs was calculated with the random effects likelihood method from Datamonkey (http: //www.datamonkey.org), and the average ratio of nonsynonymous to synonymous changes (dN/dS) was calculated using an online calculation tool (http://services.cbu.uib.no/tools/kaks). A *dN/dS* ratio above 1 implies a positive or directional selection in which advantageous mutations are being fixed, and a ratio of less than 1 implies a negative or purifying selection, suggesting the removal of deleterious mutations. Finally, quasispecies complexity was calculated using normalized Shannon entropy (Sn) as follows: $Sn = -\Sigma_i \, [p_i \cdot \ln(p_i)]/\ln(N)$, where N is the total number of sequences analyzed, and p_i is the frequency of each sequence in the

TABLE 1 Properties of sequenced data from human and pig samples

Process and/or parameter	Value by sample type		
	Human	Pig feces	Pig bile
No. of reads			
Total	28,726,064	27,146,966	25,022,058
After quality filtering	26,846,188	24,846,942	21,888,778
After host filtering	26,820,255	24,834,733	21,285,465
Matching HEV sequence	4,256	13,267	3,455,265
Assembling of HEV consensus sequences			
No. of nucleotides mapped over the HEV genome	334,924	1,035,285	269,207,053
Length of consensus HEV genome (nt)	7,261	7,259	7,317
Depth coverage (mean no. of reads per nucleotide position)	46	143	36,792
Mapping of HEV quasispecies			
No. of nucleotides mapped over the HEV genome	297,263	925,591	243,773,002
Length of consensus HEV genome (nt)	7,234	7,231	7,316
Depth coverage (mean no. of reads per nucleotide position)	41	128	33,316

viral quasispecies. Sn varies from 0 (no complexity) to 1 (maximum complexity).

Detection of nonviable mutations. In order to have an estimate of the percentage of the nonviable HEV population, mutations creating internal stop codons and mutations changing an amino acid into a proline were considered. We assumed that any stop codon within one of the three ORFs would produce a nonviable virion. Prolines are known to disrupt secondary structures and thus affect proper folding of proteins. ORF2 of HEV has been fully characterized (31) as coding for the capsid protein. Any mutations creating an additional proline in ORF2 would disrupt the structure of the capsid monomers, preventing its oligomerization and yielding nonviable virions. The range of frequency of these disruptive mutations was approximated from the highest mutation frequency observed in all sites to the sum of all frequencies, considering whether all disruptive mutations are situated on the same sequence or whether all disruptive sites are on different sequences.

Nucleotide sequence accession numbers. The consensus sequences for the full-length genomes of human HEV, swine HEV from feces, and swine HEV from bile were deposited in the GenBank under accession numbers JN906974, JN906975, and JN906976, respectively.

RESULTS

Experimental infection. Prior to inoculation, both pigs tested seronegative and negative for HEV RNA (Fig. 1). Oral inoculation of 2×10^9 copies of human HEV was successful, leading to virus excretion in both pigs from 2 dpi and seroconversion at 21 dpi in pig 2 (Fig. 1). Peak viral excretion reached 4×10^8 copies of HEV RNA/g of feces at 11 dpi. Shortly after the peak of excretion, light surgery was performed to collect the bile of the two infected animals. Bile samples of pig 1 and pig 2 contained 2×10^6 and 4×10^8 copies of HEV RNA/ml of bile, respectively. Subsequent feces samples of pig 1 and pig 2 collected at 16 dpi reached 8×10^7 and 3×10^6 copies of HEV RNA/g of feces, respectively (Fig. 1).

Generation of HEV consensus genomes from Illumina sequencing data. Illumina sequencing generated around 27×10^6 reads per sample. An average of 10% of the reads was discarded after quality filtering and mapping over the host species genomes; 0.15% to 15% of these reads matched HEV sequences (Table 1). For each of the three samples, numerous contigs were assembled, and three consensus sequences were derived with 4.3×10^3 to 3.5×10^6 reads. HEV consensus genomes differed in length (Table 1), with sequences from human feces and pig feces being shorter

than the one from pig bile of 3 nt at the 5′ untranslated region (UTR) and 53 to 55 nt at the 3′ UTR. These differences correspond to a lower coverage of the extremities due to the trimming process of each read.

The HEV consensus sequences of the three samples were 100% identical. As expected, they were found to be of genotype 3, subtype 3f. HEV genome coverage ranged from 2 to 146,597 reads per nucleotide position, depending on sample and genomic region (Fig. 2). The pool of swine bile samples had the highest HEV load: 2.56×10^9 copies of HEV RNA/ml after amplification compared to 7.42×10^4 and 6.44×10^6 copies of HEV RNA/ml for the pool of swine feces and the human sample, respectively (data not shown). As a result, the pool of swine bile sample had the highest genome coverage, with a mean coverage of 36,792 reads per nucleotide position compared to 143 and 46 reads for swine feces and human feces, respectively (Table 1).

HEV polymorphism parameters. The number of sequencing errors was plotted against the number of nucleotides mapped over conserved genes of host species and bacteria. The error rate was found to be 0.28% (Fig. 3). Because of the coverage differences, the number of polymorphic sites above the error rate found in each sample varied according to the coverage of HEV sequences in each sample: 42 SNPs for the human sample, 172 SNPs for the pool of pig feces, and 614 SNPs for the pool of pig bile, which represented 0.5%, 2.4%, and 8.3% of the genome, respectively (Table 2). Mutations occurring at a frequency as little as 1/356 could be detected, and the proportion of the HEV population displaying one particular SNP could be as high as 33% (Fig. 4).

This polymorphism is not constant along the genome (Fig. 4). The HEV quasispecies from the pool of pig bile presented values of intrahost nucleotide diversity higher for the 5′ untranslated region (UTR) and hypervariable region than for the region coding for the RdRP (1.4%, 0.093%, and 0.044%, respectively) (data not shown).

Mean diversity ranged from 0.03% for the human sample to 0.18% for the pool of pig bile, whereas intrahost nucleotide diversity ranged from 0.028% to 0.07%. The type of mutations found in the three samples gave an unusual rate of transition/transversion of around 0.6. From 50 to 85% of SNPs resulted in nonsynonymous mutations, which represented 1.4% to 14.1% of the total
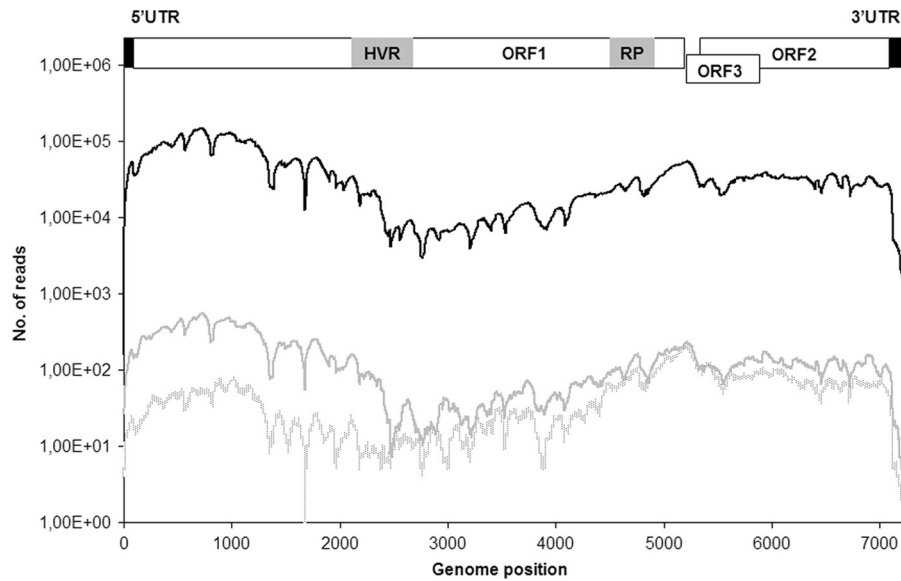
FIG 2 Coverage of the HEV genome by high-throughput sequencing. A schematic representation of the HEV genome is shown at the top of the figure. ORFs are drawn to scale, and the UTR, the hypervariable region (HVR), and the region coding for the RNA polymerase (RP) are highlighted. Numbers of reads are projected along the genomic position. Black line, bile samples pooled from the two pigs; plain gray line, feces sample pooled from the two pigs; dotted gray line, human sample.

length of the three combined ORFs (Table 2). Selective pressure along the three ORFs was mainly neutral, with a few negatively selected sites in ORF2; no positively selected sites ($dN > dS$) could be detected (data not shown). The average genome-wide dN/dS ratio ranged from 0.91 to 0.51 (Table 2), suggestive of a negative selection. Finally, genome-wide normalized Shannon entropy was fairly low, ranging from 0.006 to 0.011 (Table 2).

A total of 5 to 32 sites with mutations creating internal stop codons or additional prolines could be detected in the HEV sequences of the three samples. The frequency of these disruptive mutations in the HEV population ranged from 2.7 to 20.7% (Table 2).

Conserved polymorphism. Ninety-two SNPs were shared by the sequences from the pool of pig bile or feces, but, more importantly, 22 polymorphic nucleotide positions were shared by the sequences from human and pig bile; and 12 SNPs were shared by the sequences of all three samples (Fig. 4 and Table 3). Of these 12 SNPs, 6 were situated in ORF1, 3 were in the overlapping fragment of ORF2 and ORF3, and 3 others were in ORF2 alone. Only two of these mutations were transitions, and only two resulted in synonymous amino acid changes. The average frequency of these 12 shared SNPs were, respectively, 1.6, 3.5, and 4 times higher than the average SNP frequency in the human sample, the pig feces, and the pig bile samples (Tables 2 and 3).
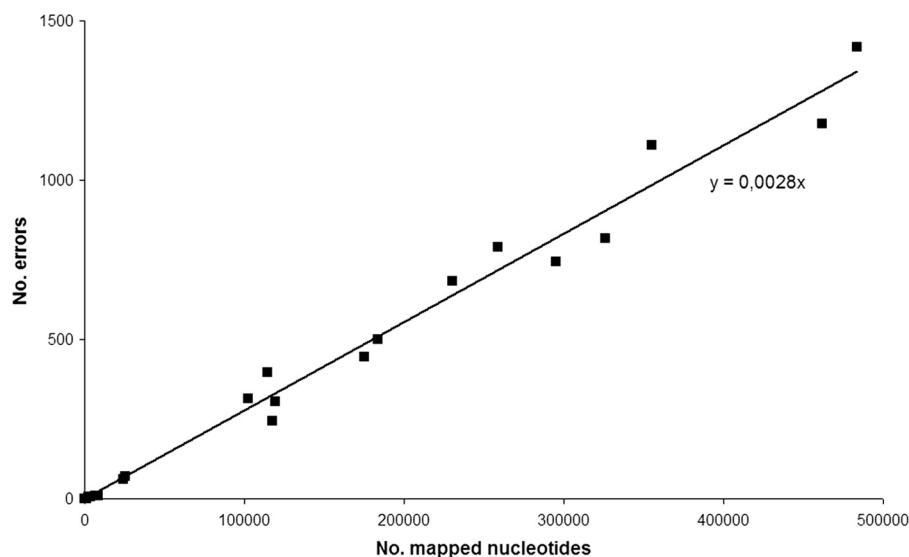

FIG 3 Plot of the number of sequencing errors versus the number of nucleotides mapped over conserved genes of host species and bacteria from our HEV samples. The black line represents the linear regression resulting in an error rate of 0.28%.

TABLE 2 Statistics on polymorphism of HEV from human and pig samples

Sample source	No. of SNPs (% of the genome)	Avg SNP frequency (%)	Mean diversity (%)[a]	π (%)[b]	Ts/Tv[c]	No. of variable amino acid sites (%)	dN/dS[d]	Sn[e]	No. of sites with deleterious mutations[f]	Frequency of deleterious mutations (%)
Human	42 (0.5)	4.93	0.03	0.028	0.35	36 (1.4)	0.91	0.006	5	7-19.6
Pig feces	172 (2.4)	1.17	0.19	0.047	0.67	110 (4.3)	0.72	0.008	6	8.1-17
Pig bile	614 (8.3)	0.83	0.18	0.070	0.74	355 (14.1)	0.51	0.011	32	2.7-20.7

[a] Calculated as the percentage of the number of substitutions divided by the total number of nucleotides.
[b] π is the nucleotide diversity calculated as the average percentage of SNPs over the genome (27).
[c] Transition/transversion ratio.
[d] Ratio of nonsynonymous to synonymous mutations.
[e] Shannon entropy.
[f] Internal stop codons and disruptive proline.

Comparison of polymorphism with other viruses. The values for intrahost nucleotide diversity (π) are consistent with those of other viruses (Fig. 5). The average π of the full-length HEV genome varied according to the coverage of HEV sequences in each sample and ranged from 0.028% in the human sample to 0.07% in the pool of pig bile. These values are in the range of the values obtained for zoonotic viruses, such as the West Nile virus (WNV) (0.021% in birds to 0.034% in mosquitoes) (15), but are in the lower range of viruses present in humans, such as the human immunodeficiency virus (HIV) (range, 0.04 to 2.5%), or even five times lower than values found for hepatitis C virus (HCV) (range, 0.04 to 4.1%) (33, 34).

DISCUSSION

Zoonotic transmissions between humans and swine have been highly suspected since partial HEV sequences from both hosts can share more than 99% identity (4) and since experimental cross-infection of subtype 3a HEV in pigs and primates leads to productive HEV infections (24). Subtype 3f HEV is the most common subtype in France and Europe (4, 22, 40) and has been shown to be circulating actively between humans and swine (4). This subtype was selected to study the effect of an interspecies transmission on the genomic adaptation of HEV in its full-length consensus sequence and its quasispecies.

The oral route mimics natural infection of this enterically transmitted disease but has been shown to be less efficient than intravenous or intrahepatic routes (18). In the present study, oral exposure of pigs to human subtype 3f HEV led to a productive HEV infection. A previous observation that oral exposure to human genotype 1 did not give rise to infection may thus have been related to the restriction of genotype 1 to humans rather than to the route of inoculation (2, 21).

Surprisingly, no nucleotide mutations could be found over the full-length consensus sequence amplified after the interspecies transmission, which demonstrates a clear adaptation of genotype 3 HEV to both humans and swine. Additionally, 29% (12/42) of polymorphic sites of HEV from the human sample were effectively infectious and found to be excreted in the feces of pigs at 16 dpi. As demonstrated in other studies, this spectrum of mutations does not necessarily increase fitness of one virion but might rather lead to increased infectivity and zoonotic potential through the diversity of HEV quasispecies population (7, 39).

Conversely, not all polymorphic sites could be transmitted since a large number of mutations were deleterious. At least 2 to 20.7% of HEV quasispecies population have been found to be nonviable. These results represent the lowest range of nonviable sequences since no mutations other than stop codons or proline were considered. Sanjuan et al. estimated that up to 40% of random mutations in RNA viruses are lethal (35).

The ratio of transitions/transversions observed in the present study is indicative of whether or not the mutations observed are random. In phylogeny, a bias toward a ratio of 1 is commonly observed since transitions seem favored over transversions, possibly as a result of the underlying chemistry of mutation. In the present study, a ratio closer to 0.5 has been observed, suggesting that mutations seem to occur at random. It is then possible to infer that the development of HEV quasispecies occurs at random, resulting in a high proportion of deleterious mutants, as stated by Sanjuan et al. (35). HEV quasispecies is then purified of its deleterious mutations, as shown by the negative dN/dS ratio.

As Belshaw et al. discussed, mutations and substitutions occur at different tempos and at different biological levels (3). Substitutions are defined as mutations which are fixed in a population. The present study dealt with nonfixed mutations undetectable at the level of the consensus sequence but observed at the level of the quasispecies and therefore expressed as the average percentage of SNPs. Nonetheless, a correlation in the variation of the mutation rate along the genome observed in this study could be made with a previous report studying the substitution rate along HEV genomes. Variation in the substitution rate along the genome of HEV has been predicted previously as being lower for the region encoding the RdRP (8.4×10^{-4} substitutions per site per year) than for the complete genome (1.51×10^{-3} substitutions per site per year) (30). In the present study, the mutation rate was also observed to be significantly lower in the RdRP (0.044%) than in other parts of the genome (up to 1.4%), which may be explained by higher functional constraints on this coding region.

In the end, HEV quasispecies resulted in a low-diversity and low-complexity population compared to other human RNA viruses such as HIV or HCV (33, 34). HEV intrahost nucleotide diversity is closer to what has been found for the zoonotic virus WNV. Indeed, viruses that need to infect diverse hosts to produce a full viral cycle, like arboviruses, are subjected to higher constraints and thus evolve more slowly than other RNA viruses (15).

In addition to being a useful tool for discovering new pathogens (5, 36, 37), HTS is also of great interest in delineating the quasispecies of viruses since the use of specific PCRs to amplify subgenomic regions of the virus, which could introduces bias, is avoided. But great care should be put into the handling of poly-
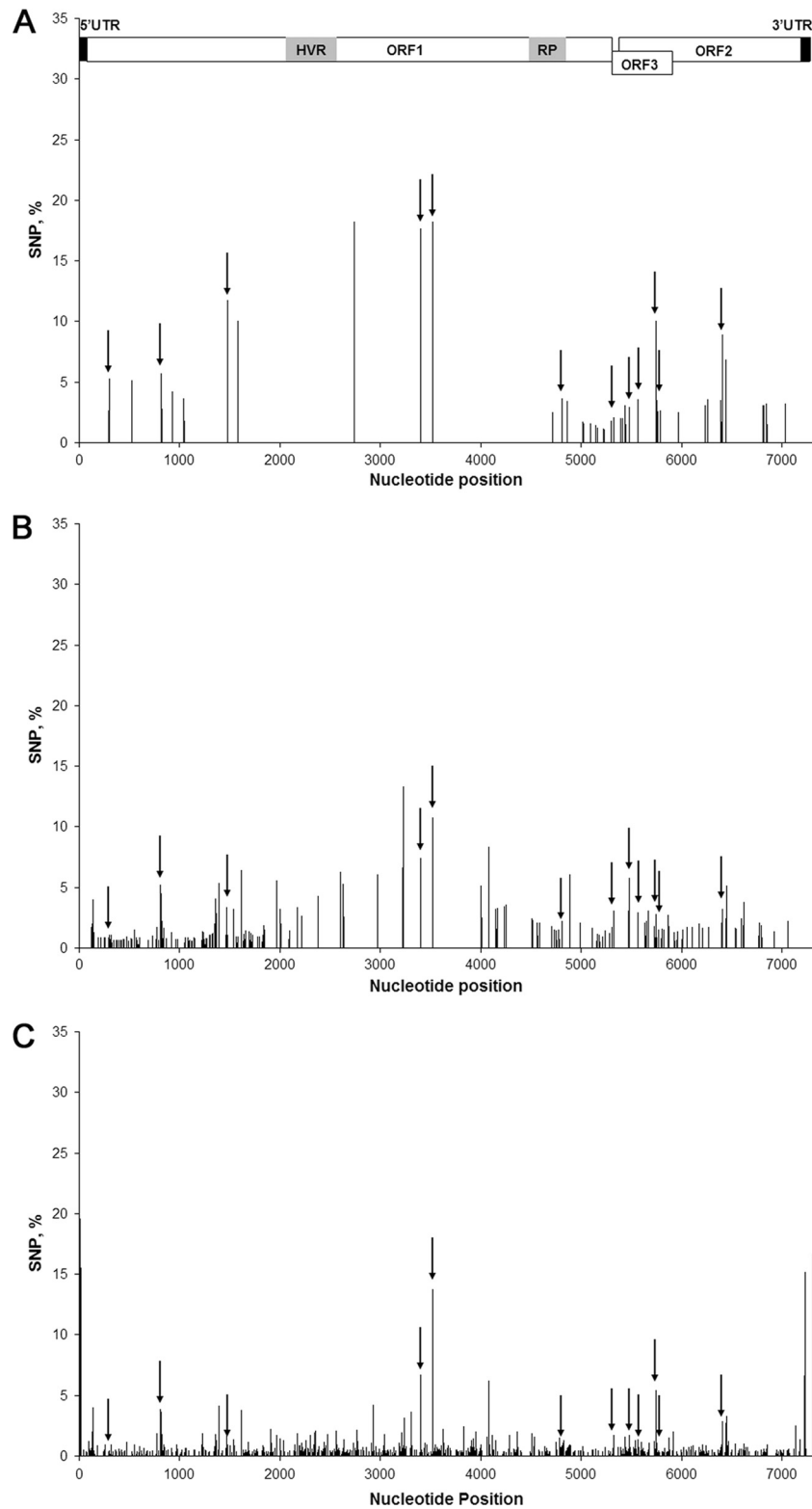
FIG 4 SNPs along the HEV genome. A schematic representation of the HEV genome is aligned at the top of the figure (see the legend of Fig. 2). Bars indicate the percentage of validated SNPs along the HEV genome sequenced from the human sample (A), the feces of pigs (B), and the bile of pigs (C). Black arrows indicate the position of SNPs shared by the three samples.

TABLE 3 Single nucleotide positions shared by HEV sequences from human and pig

Consensus sequence			Human			Pig feces			Pig bile		
Position (nt) and/or ORF	Nucleotide	Amino acid	SNP (%)	Nucleotide	Amino acid	SNP (%)	Nucleotide	Amino acid	SNP (%)	Nucleotide	Amino acid
297	C	A	5.26	G	C	1.05	G/A	C/D	0.34	G	C
820	A	L	5.71	C	F	4.50	C/T	F/F	3.60	C	F
1474	T	C	11.76	G	W	1.04	G	W	0.91	G	W
3404	C	H	17.65	T	H	7.41	T	H	6.70	T	H
3519	T	V	18.18	G	G	10.71	G	G	13.72	G	G
4809	T	V	3.64	G	G	2.22	G	G	0.88	G	G
5325[a]	G		2.06	T/C		3.03	T/C		1.74	T/C	
ORF3		R			R/R			R/R			R/R
ORF2		A			S/P			S/P			S/P
5481[a]	A		2.94	C		5.75	C		1.73	C	
ORF3		E			D			D			D
ORF2		T			P			P			P
5568[a]	A		3.57	C		2.94	G		1.37	C	
ORF3		L			L			L			L
ORF2		T			P			A			P
5743	A	N	10.00	C	T	2.75	C	T	5.43	C	T
5749	T	L	7.04	C	P	1.80	C	P	1.57	C	P
6407	G	G	8.89	C	G	3.23	C	G	2.82	C	G/G

[a] Position situated in the overlapping region of ORF2 and ORF3.

morphic data since various biases are reported for HTS techniques ([13], [26]).

Sequence coverage depth is, for example, critical when different samples are compared. Variations of polymorphism observed in this study between the three samples should not be considered as properties of HEV quasispecies in different hosts or sampling points but as a consequence of the differences in nucleotide coverage. A higher sequence coverage contains more information and is therefore more accurate in detecting low-frequency variants. The lower number of SNPs and the smaller intrahost nucleotide diversity observed for HEV from the human sample than from the pig samples are only the results of its lower coverage.

Interestingly, the mutation error rate for Illumina GaIIx calculated in this study as being 0.28% was the same as previously reported ([26]). A number of insertions/deletions were found in the HEV sequences, all of which fell under the mutation error rate. The insertion/deletion rate generated by the amplification and high-throughput sequencing processes could be calculated as being 5.7×10^{-7} (data not shown), which is lower than what has been previously established (4×10^{-6}) ([26]). This insertion/deletion rate was very likely reduced by the second mapping of HEV sequences, which removed all reads containing more than two mismatches.

Here is presented the first report on the use of HTS for the study of full-length genomes of HEV and, more generally, on the use of HTS to analyze viral variability upon interspecies transmission. The observation that the full-length consensus sequence of HEV is conserved in spite of a change of host demonstrates the absence of a species barrier and the clear adaptation of genotype 3f HEV to both hosts. Moreover, this study confirms that HEV exists as a quasispecies in the *in vivo* setting and that genetic variability extends throughout its genome. Finally, major SNPs were conserved during the interspecies transmission. These results may suggest that transmission of swine HEV to humans would result in the absence of adaptation and in a productive HEV infection.

In conclusion, the transmission of human HEV to pigs did not seem associated with a restriction in genetic diversity, most likely because HEV infection of either host does not impact its viral cycle. According to the typology of zoonosis proposed by Pepin et al. ([29]), the transmission of some zoonotic agents can be governed only by ecological drivers. In this case, all viral genotypes circulating in the reservoir are already competent for transmission in the new host. Founder effects or adaptative fine-tuning in the new host could explain the variability of the strains. These results suggest that HEV could belong to this category of viruses.
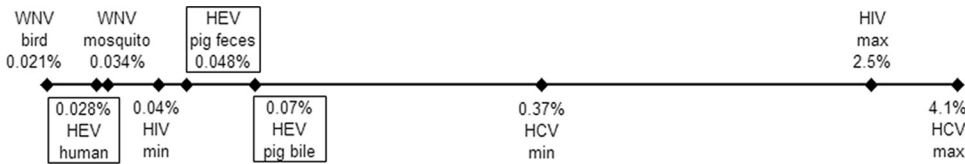

FIG 5 Viral intrahost nucleotide diversity (π) of viruses. Values of π are displayed as percentages on a logarithmic number line for HEV alongside the West Nile virus (WNV), human immunodeficiency virus (HIV), and hepatitis C virus (HCV) ([15], [33], [34]).

ACKNOWLEDGMENTS

REFERENCES

1. **Arankalle VA, Chobe LP, Chadha MS.** 2006. Type-IV Indian swine HEV infects rhesus monkeys. J. Viral Hepat. **13**:742–745.
2. **Balayan MS, Usmanov RK, Zamyatina NA, Djumalieva DI, Karas FR.** 1990. Brief report: experimental hepatitis E infection in domestic pigs. J. Med. Virol. **32**:58–59.
3. **Belshaw R, Sanjuán R, Pybus OG.** 2011. Viral mutation and substitution: units and levels. Curr. Opin. Virol. **1**:430–435.
4. **Bouquet J, et al.** 2011. Close similarity between sequences of hepatitis E virus recovered from humans and swine, France, 2008–2009. Emerg. Infect. Dis. **17**:2018–2025.
5. **Buckwalter MR, et al.** 2011. Identification of a novel neuropathogenic Theiler's murine encephalomyelitis virus. J. Virol. **85**:6893–6905.
6. **Cheval J, et al.** 2011. Evaluation of high-throughput sequencing for identifying known and unknown viruses in biological samples. J. Clin. Microbiol. **49**:3268–3275.
7. **Coffey LL, Vignuzzi M.** 2011. Host alternation of chikungunya virus increases fitness while restricting population diversity and adaptability to novel selective pressures. J. Virol. **85**:1025–1035.
8. **Feagins AR, Opriessnig T, Huang YW, Halbur PG, Meng XJ.** 2008. Cross-species infection of specific-pathogen-free pigs by a genotype 4 strain of human hepatitis E virus. J. Med. Virol. **80**:1379–1386.
9. **Graff J, et al.** 2005. The open reading frame 3 gene of hepatitis E virus contains a *cis*-reactive element and encodes a protein required for infection of macaques. J. Virol. **79**:6680–6689.
10. **Grandadam M, et al.** 2004. Evidence for hepatitis E virus quasispecies. J. Gen. Virol. **85**:3189–3194.
11. **Halbur PG, et al.** 2001. Comparative pathogenesis of infection of pigs with hepatitis E viruses recovered from a pig and a human. J. Clin. Microbiol. **39**:918–923.
12. **Haqshenas G, Shivaprasad HL, Woolcock PR, Read DH, Meng XJ.** 2001. Genetic identification and characterization of a novel virus related to human hepatitis E virus from chickens with hepatitis-splenomegaly syndrome in the United States. J. Gen. Virol. **82**:2449–2462.
13. **Harismendy O, et al.** 2009. Evaluation of next generation sequencing platforms for population targeted sequencing studies. Genome Biol. **10**: R32. http://genomebiology.com/content/10/3/R32.
14. **Huang CC, et al.** 1992. Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HEV). Virology **191**:550–558.
15. **Jerzak G, Bernard KA, Kramer LD, Ebel GD.** 2005. Genetic variation in West Nile virus from naturally infected mosquitoes and birds suggests quasispecies structure and strong purifying selection. J. Gen. Virol. **86**: 2175–2183.
16. **Johne R, et al.** 2010. Novel hepatitis E virus genotype in Norway rats, Germany. Emerg. Infect. Dis. **16**:1452–1455.
17. **Jothikumar N, Cromeans TL, Robertson BH, Meng XJ, Hill VR.** 2006. A broadly reactive one-step real-time RT-PCR assay for rapid and sensitive detection of hepatitis E virus. J. Virol. Methods **131**:65–71.
18. **Kasorndorkbua C, et al.** 2004. Routes of transmission of swine hepatitis E virus in pigs. J. Clin. Microbiol. **42**:5047–5052.
19. **Koonin EV, et al.** 1992. Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus: delineation of an additional group of positive-strand RNA plant and animal viruses. Proc. Natl. Acad. Sci. U. S. A. **89**:8259–8263.
20. **Li T-C, et al.** 2005. Hepatitis E virus transmission from wild boar meat. Emerg. Infect. Dis. **11**:1958–1960.
21. **Lu L, et al.** 2004. Complete sequence of a Kyrgyzstan swine hepatitis E virus (HEV) isolated from a piglet thought to be experimentally infected with human HEV. J. Med. Virol. **74**:556–562.
22. **Lu L, Li C, Hagedorn CH.** 2006. Phylogenetic analysis of global hepatitis E virus sequences: genetic diversity, subtypes and zoonosis. Rev. Med. Virol. **16**:5–36.
23. **Meng XJ, et al.** 1998. Experimental infection of pigs with the newly identified swine hepatitis E virus (swine HEV), but not with human strains of HEV. Arch. Virol. **143**:1405–1415.
24. **Meng XJ, et al.** 1998. Genetic and experimental evidence for cross-species infection by swine hepatitis E virus. J. Virol. **72**:9714–9721.
25. **Meng XJ, et al.** 2011. *Hepeviridae*, p 991–998. *In* King AMQ, et al (ed), Virus taxonomy. Ninth report of the International Committee on Taxonomy of Viruses. Elsevier, London, United Kingdom.
26. **Minoche AE, Dohm JC, Himmelbauer H.** 2011. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems. Genome Biol. **12**:R112. http://genomebiology.com/content/12/11/R112.
27. **Nei M, Li WH.** 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. U. S. A. **76**: 5269–5273.
28. **Pavio N, Meng X-J, Renou C.** 2010. Zoonotic hepatitis E: animal reservoirs and emerging risks. Vet. Res. **41**:46.
29. **Pepin KM, Lass S, Pulliam JRC, Read AF, Lloyd-Smith JO.** 2010. Identifying genetic markers of adaptation for surveillance of viral host jumps. Nat. Rev. Microbiol. **8**:802–813.
30. **Purdy MA, Khudyakov YE.** 2010. Evolutionary history and population dynamics of hepatitis e virus. PLoS One **5**:e14376.
31. **Robinson RA, et al.** 1998. Structural characterization of recombinant hepatitis E virus ORF2 proteins in baculovirus-infected insect cells. Protein Expr. Purif. **12**:75–84.
32. **Rose N, et al.** 2011. High prevalence of Hepatitis E virus in French domestic pigs. Comp. Immunol. Microbiol. Infect. Dis. **34**:419–427.
33. **Sakai A, Kaneko S, Honda M, Matsushita E, Kobayashi K.** 1999. Quasispecies of hepatitis C virus in serum and in three different parts of the liver of patients with chronic hepatitis. Hepatology **30**:556–561.
34. **Salazar-Gonzalez JF, et al.** 2009. Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. J. Exp. Med. **206**:1273–1289.
35. **Sanjuán R, Moya A, Elena SF.** 2004. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proc. Natl. Acad. Sci. U. S. A. **101**:8396–8401.
36. **Sauvage V, et al.** 2011. Identification of the first human gyrovirus, a virus related to chicken anemia virus. J. Virol. **85**:7948–7950.
37. **Sauvage V, et al.** 2011. Human polyomavirus related to African green monkey lymphotropic polyomavirus. Emerg. Infect. Dis. **17**:1364–1370.
38. **Tei S, Kitajima N, Takahashi K, Mishiro S.** 2003. Zoonotic transmission of hepatitis E virus from deer to human beings. Lancet **362**:371–373.
39. **Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R.** 2006. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature **439**:344–348.
40. **Widén F, et al.** 2011. Molecular epidemiology of hepatitis E virus in humans, pigs and wild boars in Sweden. Epidemiol. Infect. **139**:361–371.